

PROCESSAMENTO DE FALA 2013/14

2º Teste

23 de Maio de 2014

Nome:	
Número:	
Número do grupo:	

Indique símbolos fonéticos usando o alfabeto SAMPA (dialecto de Lisboa).
As respostas erradas às perguntas V/F serão penalizadas.

Cotação: 2,0 + 1,6 + 1,0 + 1,0 + 1,5 + 1,0 + 1,5 + 0,8 + 1,6 + 1,8 + 1,0 + 2,6 + 2,6

1. Considere um codificador RPE-LTP a 14kbsp. A qual das partes, relativas aos nomes RPE ou LTP são atribuídos mais bits. Justifique sucintamente, indicando **apenas** os cálculos do débito em kbps que atribuiria a cada uma destas duas partes.

--

2. Indique dois codificadores de cada classe:

- (a) Codificadores de forma de onda no domínio da frequência
- (b) Codificadores análise por síntese
- (c) Vocoders
- (d) Codificadores que modelam explicitamente o carácter misto da excitação nas várias subbandas

a	
b	
c	
d	

3. Indique dois testes que tanto possam ser usados na avaliação de codificadores como de sintetizadores.

--

4. Como se designam (indique o nome completo em inglês):

- (a) o modelo de anotação prosódica com várias fiadas?

(b) o modelo de síntese por concatenação com menor custo computacional?

a	
b	

5. Diga se as seguintes afirmações são verdadeiras (V) ou falsas (F), colocando V/F à esquerda da alínea:

- (a) Numa língua tonal, a entoação pode distinguir semanticamente as palavras.
- (b) Em síntese de fala, as regras de duração derivadas quer manualmente quer automaticamente contemplam um factor atenuador para vogais em final de frase.
- (c) Na síntese por concatenação, o corpus de fala deve ser sempre lido mantendo uma entoação o mais uniforme possível.
- (d) O modelo de excitação de massas e molas é frequentemente usado em sintetizadores de formantes.
- (e) A distância entre coeficientes cepstrais e a distância entre valores de F0 pesam igualmente no cálculo de custos de concatenação de duas unidades.

6. Considere a síntese por concatenação de difones de uma fricativa genericamente designada por F. O módulo de geração prosódica determina para a duração de F o valor alvo de 120 ms. No inventário de unidades, dispõe-se do difone VF (com valores de duração 90 e 40 para as duas sub-unidades) e do difone FV (com valores de duração 60 e 80 para as duas sub-unidades), em que V designa genericamente uma vogal. Indique o factor multiplicador que deve afectar as duas partes de F a concatenar.

--

7. Considere que tem um corpus de um orador dizendo os dias da semana (domingo, segunda-feira, ... sábado), utilizado num sintetizador de domínio limitado. Indique 7 palavras alternativas que esse sintetizador por concatenação poderia construir garantindo todas as transições entre fones (palavras que não garantam isso serão penalizadas). Utilize o símbolo # para silêncio. Pode considerar que o acento secundário e primário são equivalentes para este exercício.

--

8. Faça a transcrição fonética das palavras estrangeiras *Championship Manager*, contemplando duas pronúncias alternativas, em que uma delas (transc1) é a derivada com as regras do léxico comum, e outra (transc2) a pronúncia mais corrente nos media.

transc1	
transc2	

Nome:	
Número:	
Número do grupo:	

9. Utilizando a seguinte sintaxe

$a \rightarrow b / c \dots d$

descreva as duas regras que dão conta da pronúncia do ditongo final dos verbos regulares da primeira conjugação, presente do indicativo, terceira pessoa, plural. c e d podem ser grafemas (ex: a, b, etc.), classes de grafemas (vogais, consoantes, etc.), a fronteira de palavra (#), ou qualquer grafema (*). Pode usar o símbolo 0 para marcar nulos fonémicos, e os símbolos | e () para marcar disjunção entre vários grafemas (ex: a | b | c). As regras são aplicadas por ordem até que uma emparelhe e, nesse caso, já não se aplicam as seguintes. (As regras não precisam de contemplar palavras compostas.)

--

10. Diga se as seguintes afirmações são verdadeiras (V) ou falsas (F) , colocando V/F à esquerda da alínea:

- (a) A utilização de modelos de misturas de Gaussianas evita a degradação inerente à quantificação vectorial de coeficientes cepstrais.
- (b) A interpolação de modelos de língua é uma alternativa para adaptar os reconhecedores a novos domínios, utilizando um subconjunto para calcular os pesos da interpolação.
- (c) O número de estados de cada palavra um sistema de reconhecimento de palavras isoladas baseado em HMMs é tipicamente igual ao número de sílabas.
- (d) Na estrutura de uma sílaba, a coda é um elemento mais estável que o ataque em termos de variantes de pronúncia.
- (e) O algoritmo de Viterbi é usado na resolução de dois dos problemas clássicos em HMMs.
- (f) A perplexidade aumenta linearmente com a entropia.

11. Como se designam (indique o nome completo em inglês):

- (a) o método de reconhecimento de padrões que permite comparar padrões de referência e de teste com durações diferentes?
- (b) o processo de re-treino de modelos HMM de fones com base em alinhamento forçado?

a	
b	

12. Considere um sistema de reconhecimento de fala contínua e vocabulário extenso, independente do orador. Ignorando a pontuação, complete os valores de H ("correct"), D ("deletions") , S ("substitutions"), I ("insertions"), N ("total"), %Corr, %Acc e %WER correspondentes às transcrições manuais e automáticas apresentadas, indicando apenas os cálculos. Diga qual a maior causa de erros:

- (a) presença de OOVs
- (b) presença de disfluências
- (c) dificuldade em reconhecer palavras funcionais
- (d) dificuldade em reconhecer palavras de contexto

Manual: *Não pode haver nenhum código de ética que se sobreponha àquilo que é o dever, a obrigação do médico, de ajudar os doentes, denunciar aquilo que está mal.*

O bastonário da Ordem dos Médicos fala em censura.

Mais do que %pp a indignação dos profissionais de saúde, isto deveria gerar a revolta de toda a população contra o ministério que quer impor a ro-, a lei da rolha aos profissionais de saúde e impedi-los de continuarem a defender os doentes.

Automática: *Não pode haver nenhum código de ética se sobrepõem aquilo que ele deve obrigação do médico. Ajudar os doentes de enunciar aquilo que está mal.*

O bastonário da Ordem dos Médicos está em censura.

Mais do que e a indignação dos profissionais sobre Didi isto deveria gerar revolta de toda a população contra o ministério. Querem pôr a roupa a lei da rolha aos profissionais de saúde impedi-los continuarem a defender os doentes.

Nota: Na transcrição manual foram usadas algumas convenções de anotação (%pp significa “pausa preenchida” e os fragmentos de palavras terminam em hífen).

H	D	S	I	N	% Corr	% Acc	% WER	a/b/c/d

13. Considere o corpus de treino composto pelas seguintes frases:

A Espanha venceu o campeonato de dois mil e dez.

A Itália derrotou a Alemanha em dois mil e quatro.

Portugal nunca venceu o mundial.

O mundial de dois mil e catorze é no Brasil.

O Brasil e a França já venceram o mundial.

E as frases de teste:

A Alemanha venceu o mundial.

O Brasil venceu o mundial.

A Espanha venceu o mundial.

A França venceu o mundial.

A Itália venceu o mundial.

Portugal venceu o mundial.

- Considerando um modelo de bigramas sem alisamento, a que país(es) se refere(m) a(s) frase(s) de teste com maior probabilidade?
- Calcule a probabilidade da frase de teste sobre Portugal com um modelo de bigramas com alisamento *add-one*, indicando apenas os cálculos.
- Construa uma frase com um número infinito de palavras e probabilidade não nula, de acordo com um modelo de bigramas.

a	
b	
c	